Pre-print copy of submission accepted at 12th International Symposium on Empirical Software Engineering and Measurement (ESEM), Oulu, Finland, 2018

An Empirical Investigation of Transferring Research to Software Technology Innovation

A Case of Data-Driven National Security Software

Mansooreh Zahedi
IT University of Copenhagen
Denmark
mzah@itu.dk

Muhammad Ali Babar
University of Adelaide
Australia
ali.babar@adelaide.edu.au

Brenton Cooper
Data to Decisions CRC
Australia
Brenton.cooper@d2dcrc.com.au

ABSTRACT

Context: Governments are providing more and more support for academia-industry collaborations for industry led research and innovation via Cooperative Research Centers (CRC). It is important to understand the processes and practices of such programs for transferring scientific R&D to innovation.

Goal: We aimed at empirically investigating the processes and practices implemented in the context of one of the Australian CRCs, aimed at transferring big data research to innovative software solutions for national security.

Method: We applied case study method and collected and analyzed data from 17 interviews and observations of the participants of the studied CRC program.

Findings: We present the innovation process implemented in the studied CRC. We particularly highlight the practices used to involve end-users in the innovation process. We further elaborate on the challenges of running this collaborative model for software technology innovation.

CCS CONCEPTS

• **Software and its engineering** → Software creation and management

KEYWORDS

Qualitative method, software innovation, academia-industry, technology transfer, Cooperative Research Center (CRC)

1. INTRODUCTION

The emergence of the critical role and appliance of data science in diverse domains has more than ever highlighted the necessity of academia-industry collaboration to provide innovative solutions based on the state of the art knowledge. Practitioners need access to the most recent scientific advances underpinning the solutions to the real-world problems. Researchers need business domain knowledge and access to real-world scenarios and data to design and evaluate innovative solutions. Academia-industry collaboration has been an important topic in Software Engineering (SE). There have been numerous studies to enable and sustain academia-industry collaboration (e.g., [1], [2], [3]). Yet, it is a challenging undertaking [4], [5] as researchers and practitioners differ in several areas such as preferences, goals, skillsets, problem-solving approaches, and time frames. These differences introduce complexities in communication, coordination and collaboration of the involved parties. Garousi et al. [4] have identified several challenges of academia-industry collaboration such as: lack of interest or commitment, lack of research relevance for practice and lack of trust. Studies report that a successful collaboration needs *continuous* interaction and engagement [3] supported by mechanisms such as adopting agile methodologies [1], [5].

In order to enable academia-industry collaboration, some governments (e.g., Australia) have been investing in programs like Cooperative Research Center (CRC). The goal of the CRC programs is to bring researchers, industry and the community together to apply scientific research into innovative practices and technologies [6]. There have been several reports about pros and cons of CRC programs in Australia [7]. However, there has been little empirical effort aimed at investigating the process and practices of a CRC collaboration model.

We report a case study that has studied one of the Australian CRCs for setting up and sustaining a *user-centric* collaboration between researchers and practitioners. The studied CRC was aimed at conducting world-class data science research and developing innovative data analytical software solutions for the Australian law enforcement agencies. The main contributions of this study are:

- It reports one of the first efforts aimed at empirically exploring the processes and practices of an Australian CRC program for conducting industry-driven research and transferring the outputs to software technologies.
- It reports an implementation of academia-industry collaboration, which emphasizes involving users in the process of software technology innovation.
- It highlights the challenges characterizing three-ways collaborations involving researchers-practitioners-users in R&D projects. The identified challenges point out the areas for future research.

2. BACKGROUND

We discuss the concepts and studies related to our work.

2.1. Software Technology Innovation

The concept of innovation refers to "a process that begins with an invention, proceeds with the development of the invention, and results in the introduction of a new product, process or service to the market" [8]. Garcia and Calantone [9] emphasize that the innovation process not only includes technological development of an invention, but also diffusion of that invention to end-users. They highlight that the degree of 'newness' indicates the innovativeness of a product, which can be viewed from different perspectives (e.g., new to market, to firm, to customers) [9]. In the context of Software Engineering (SE), an innovation may mean product innovation (e.g., creating new technology), process innovation (e.g., introducing new design method) or organization innovation (e.g., implementing new organizational method) [10].

Punter et al. [11] consider software technology innovation as a process of technology creation and transfer. They suggest an incremental evolution of technology creation phase, which overlaps with technology transfer on evaluation activity with an integrated feedback loop [11]. Gorschek et al. [2] propose a technology transfer process for design and adoption of innovative software technology. Diebold et al. [12] state that SE research projects' outputs are mostly not mature enough to be applied in practice. There needs to be human-intensive effort for transferring SE research to technology [12].

R&D firms are getting interested in the open innovation paradigm where they search for ideas beyond their organizations, as well as using the external channels to leverage their internal ideas [13],[14]. This trend results in forming an innovation ecosystem, which includes a network of stakeholders (e.g., researchers, developers, and customers) that are engaged in the innovation process while sharing costs and risks [15]. The new innovation models of software companies focus on involving customers in a collaborative approach following the principles of fail fast [15].

2.2. User Involvement in Innovation Process

Different approaches (e.g., participatory design and ethnography) have been employed to promote users' engagements in design and development of (high-tech) products. Until recently, user involvement in SE projects has been mainly manifested in agile methodology and user-centric design [16]. However, an increasing number of companies have started involving users in application development and enhancement. Availability of new tools and techniques has enabled many enterprises (e.g., Facebook and Microsoft) to conduct continuous experimentation with the released features [17],[18]. Bosch-Sijtsema and Bosch [19] observe that the nature of user-involvement practices vary in different phases of innovation, given the availability of data and engagement of customers [19]; a limited number of users are consciously involved through qualitative data collection approaches (e.g., workshops) during the pre-development phase but a larger volume of quantitative data is collected to evaluate the innovation after development [19].

2.3. Collaboration of Academia and Industry

Academia-industry collaboration is imperative for software technology innovation [20]. Such collaboration is prone to several challenges such as communication difficulties, lack of trust and commitment, mismatch between goals and skillsets, management and contractual issues [4],[21],[5]. Villani et al. [21] discuss that industry and academia experience distances in cognitive, geographical, organizational and social dimensions that lead to collaborative challenges. The challenges also stem from differences in problem solving approaches [5]. Industry focuses on finding solutions for the current problems, however, academia aims at identifying research gaps, and contributing to the relevant knowledge [5]. Several researchers have investigated these challenges from different perspectives and reported successful practices (e.g., [4], [22]). For example, a review [4] has identified more than 100 practices to support industry-academia collaboration.

2.3.1. Cooperative Research Center (CRC) Program

Aiming at facilitating and sustaining engagement between research providers (i.e., universities and research entities) and industry (i.e., private sector or government agencies), Australian government has been investing in Cooperative Research Center (CRC) programs [7], [23]. The CRC programs have proven to be beneficial to bridge institutional, geographical, social and knowledge gaps between industry and academia [21]. The CRCs also help build trust, enlarge R&D networks, and provide training and career path for researchers [7],[24]. Nevertheless, the CRCs also have some drawbacks such as administrative burden on researchers, relatively informal contractual commitments, governance and coordination of researchers with the CRC administrators and industry [24]. It is important to deploy innovative processes and practices for establishing strong bonding among a CRC's stakeholders (i.e., academic researchers, industry and customers) for transferring scientific research to technology. This aspect of a CRC program is the focus of our research.

3. RESEARCH METHOD

Case study is a suitable research method to explore a phenomenon in its real-life context, specifically when the border between the phenomenon and its context is blurred [25]. We conducted an exploratory case study to investigate our research objective, with the following research questions:

RQ1. How a user-centric approach of software technology innovation can be implemented in practice?

RQ2. What are the key challenges of running a user-centric software technology innovation?

3.1. Context

We studied an R&D program of an Australian Cooperative Research Center (CRC) called Data to Decisions (D2D). The D2D CRC is aimed at developing innovative data science methods and approaches to leverage big data for developing and deploying security related solutions for the Australian Defense and law enforcement agencies. We focused on the D2D's one large program aimed at producing software technologies to facilitate crime investigation. The program seeks to build uniform data integrated from different types of data sources owned by various agencies. It applies data science and SE for inventing sophisticated solutions in this domain. The studied program involves four different research groups (i.e., include leaders, research fellows, post-docs and PhD students) from two Australian universities.

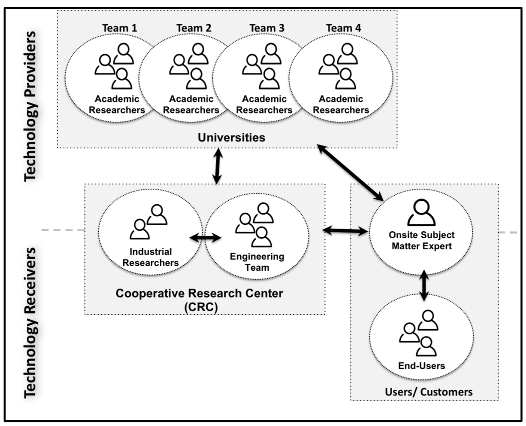

Fig 1. Stakeholders of the Studied CRC Program

The researchers contribute to the program from different angles such as inventing federated data architecture, extracting and linking similar entities in data, providing intelligent front-end system and innovative visualization of knowledge graphs. As a key player of an innovation ecosystem, the D2D collaborates with the academic researchers on the one hand and the customers and end-users on the other hand. Apart from the two senior managers, the D2D has allocated an engineering team of four developers. In order to facilitate continuous involvement of users in the process of developing science and software technologies, the D2D has involved an onsite Subject Matter Expert (SME) from the initial phases. This role facilitates communication among the involved stakeholders as shown in Fig. 1.

3.2. Data Collection and Analysis

Our data collection method included semi-structured interviews and observations[1]. We selected our interviewees from all the stakeholder groups involved in the program. We interviewed 17 participants (i.e., shown in Table 1). The majority of the interviews were conducted in person, except three interviews that were carried via Skype/phone. Each interview lasted between 40-60 minutes. All the interviews were audio recorded and verbatim transcribed. We started the interviews with senior roles (e.g., managers and research leads) from different stakeholder groups and collected the big picture of the processes and practices being followed. We further approached more interviewees from each stakeholder group, for more detailed and technical perspectives. Our interview guide was customized based on the role of interviewee and the type of the stakeholder group. For example, the interviews with group leaders uncovered researchers' intentions, goals, activities, and challenges of collaboration in the program. These views were further explored from technical perspective when we interviewed postdoc researchers. In the interviews with the software engineers, we focused on their role in the program, interactions with researchers and users, and the challenges they face. The interviews with the researchers were helpful to collect the general vision of CRC, its coordination role between stakeholders, implemented processes and practices, and the key challenges. We also observed one of the group meetings of the software engineers and researchers at the CRC office.

Table 1: Data Collection – Summary of Participants

Stakeholder Group	Role	#
Research Team1	Leader, 2 Postdocs, 1 Research Fellow	4
Research Team 2	Leader, 2 Postdocs	3
Research Team 3	Leader, 1 Postdoc	2
Research Team 4	Leader	1
CRC - Researchers	2 Managers/ Senior Researchers	2
CRC - Engineers	4 Software Engineers	4
Users/ Customers	1 Subject Matter Expert	1
Total		17

We analyzed the data using thematic analysis [26]. We performed open coding in Nvivo10 using a 4-level coding scheme (i.e., Fig. 2). At the lowest level, we had the key points extracted from data fragments, which were organized into the higher levels of abstraction forming categories and high-level themes. We coded the data segments related to phases and process of software innovation in the program, mechanism to coordinate stakeholders, and the associated challenges.

[1] Due to non-disclosure agreement we are not allowed to openly share the data.

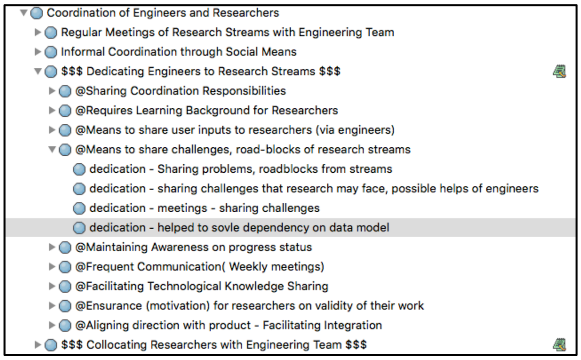

Fig. 2: A Snapshot of Coding Scheme in Nvivo

We also coded the scattered quotations elaborating the motivators for the implemented practices. We prepared and maintained extensive memos to track the emerging ideas. During the write up, we revisited all the memos, which helped in building our arguments. To verify our interpretations, we presented the findings to a selected number of the CRC stakeholders and received feedback.

4. FINDINGS

We present our findings including: the scope of innovation, the implemented innovation process, and coordination mechanisms supporting the process (RQ1) and the key challenges that the stakeholders faced in this setup (RQ2).

4.1. The Scope of Innovation

To identify the scope of innovation, we specifically asked the interviewees to reflect on what they perceived to be innovative in this program. Most of the interviewees highlighted that the software product is innovative. Because it contains state-of-the-art artifacts that are purposefully designed and implemented to solve real use cases in a very new domain (i.e., law enforcement). The requirements of the product and use cases resulted from an extensive gap analysis of law enforcement domain in Australia with potential end-users and experts. The use cases are being addressed through research and developing new knowledge and capabilities. It was stated that: *"The innovation results from the innovative conceptual nature of the artifacts that we build and the fact that we produce capabilities that could not be implemented before… the added knowledge results from the evaluations as what works and what doesn't."-**P4, Leader***. It was emphasized that innovation takes place when research has impact in practice: "*We managed to do research and [it] is going to have an impact, in the sense that it's going to lead to a product that will be used in practice. The first experiment that CRC already did hint that the investigators liked [it]."-**P16, Leader***. One of the interviewees explained the possible impact of their research with an example: *"it may turn out that juries can understand crime scenes a lot better with virtual reality, than going to the scene when it has been cleaned up after six months*"–***P10, Researcher.*** Beside the product, few interviewees stated that the implemented process through the CRC is also innovative. The process has been unique in initiation (e.g., gap analysis and defining use cases) and execution (e.g., coordination of researchers, users and engineers towards the product). It was stated that: "*We've come about defining the innovations through the process of capability-gap analysis and product roadmap […] we've defined areas, which we believe are truly innovative. […] We're applying our agile methodology […] but, one of the things I see unique is, to actually getting university researchers to work within that sort of process framework […] I think these are two areas: product innovation and process innovation."- **P6, CRC Manager.***

4.2. The Innovation Process

In this section, we elaborate on the innovation process implemented in this case study. Based on our analysis, we have identified three major phases in the process and the associated activities, as shown in Fig. 3 and elaborated below.

4.2.1. Initiation: Setting Long-term Research Vision.

This phase includes defining the problem and setting the long-term vision for research. It was driven by the CRC researchers and external consultants through extensive field studies. They interviewed almost thirty subject matter experts from defense and various law enforcement and intelligence agencies. The objective was to identify the high-priority challenges that were perceived by experts in this domain when applying analytics on users' data. The researchers identified a list of use cases and the required capabilities. The field study provided the researchers with a broader knowledge of the business domain and enabled them to identify the areas of interest to apply scientific research. The researchers conducted separate workshops on different aspects of law enforcement and intelligence with subject matter experts and potential users. Those workshops enabled the researchers to narrow down the problem into 2-3 high priority use cases. Having defined the problem area, the researchers with suitable capabilities got involved in the program through a call for participation. This process resulted in forming the research streams.

4.2.2. Execution I: Linking Research to User Needs.

One of the key activities of this phase was to maintain a continuous user focus during the program execution. It was ensured by involving a Subject Matter Expert (SME) from Australian Federal Police (AFP), who represented customers. This role remained with the program during the whole process of software technology invention and transfer. The SME reviewed the defined projects for the streams and provided feedback to align them with the users' needs. In this phase, a candidate solution of an intelligent front-end system for investigation was prototyped by one of the research streams in a close collaboration with the SME and the CRC researchers. The prototype was presented to the potential end-users for validation and collecting feedback.

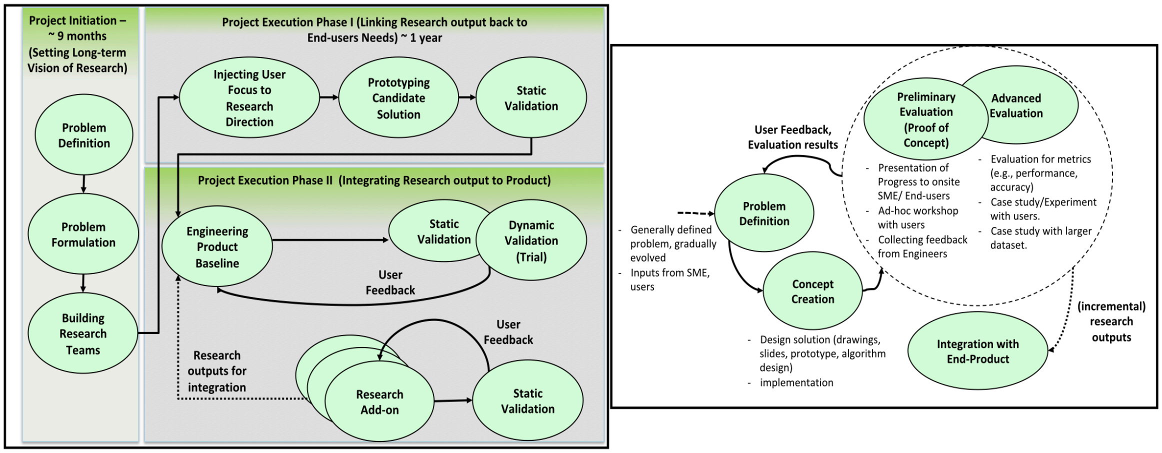

Fig. 3- A Conceptualized View of the Innovation Process: The Whole Process (left), Activities in Research Cycles (right)

4.2.3. Execution II: Integrating Research with Product. This was the longest phase in the innovation process. At the time of our study, the program had reached to this phase after running for three years. In this phase, the software engineering team got involved in the program to take over the implementation tasks from the researchers. The software engineers initiated with receiving the candidate solution, which was prototyped by researchers. They became responsible to re-engineer the prototype and develop the product baseline. We observed that since then the program has been incrementally progressing in two parallel lines: i) enhancing the product baseline and ii) progressing with the research outputs. The software engineers enhance the product based on users' feedback. They frequently interact with the SME and end-users to demonstrate the current features of the product (i.e., static validation) and apply improvements. The SME and the software engineers closely collaborated to run a couple of trials to dynamically validate the solution. The trials were conducted with end-users using the product to solve realistic scenarios. The users' feedback was collected qualitatively and quantitatively (i.e., log of application) and shared with the relevant researchers. Figure 3 (on right side) shows the key activities inside the research streams. The academic researchers adopt an incremental approach to solve the problem and provide outputs for integration in the software product. Each stream started with a general research problem, which emerged and was relatively defined in the initiation phase. They have broken down the general problems and refined them in iterations based on the research ideas and the feedback received from the onsite SME and the users on their conceptual solutions. The conceptual solutions differ in granularity varying from drawings to algorithms, software architecture and implemented prototypes. The output of each stream goes to several rounds of (preliminary) evaluation for the proof of concept. The research outcomes and the software solutions are presented to the stakeholders (i.e., CRC, SME and users) every three months, in addition to interaction with the onsite SME and the engineering team on weekly basis. Some of the research streams further conduct advanced evaluation (e.g., performance and accuracy) with the users' data when their proposed solution becomes more mature. We observed that the engineers and researchers actively engaged. They pull research outputs and integrate them into a product, whenever there is a mature output that is approved by the onsite SME in particular, and the end-users. The integrated output goes through further evaluation with users.

4.3. Coordination Mechanisms

Our analysis has revealed several coordination mechanisms implemented to support collaboration among the involved parties. Table 2 presents an overview of these mechanisms in relation to their role to coordinate between different groups of stakeholders. We elaborate the key mechanisms that were helpful to incorporate users' inputs in the innovation process.

Table 2: Coordination Mechanisms between Stakeholders

Stakeholder Groups	Coordination Mechanisms
Engineer-Researcher	Collocating Researchers with Engineers
	Dedicated engineer to each research stream & Weekly meetings
	Informally via social means, online tools
	Via CRC/ Road-map meetings
Engineer-User	Direct interaction on product
	Workshops/meetings with onsite SME
Researcher-User	Via Onsite SME (regularly)
	Via Engineering Team
	Via CRC/ Road-map meetings
	Via CRC/ ad-hoc Workshops with users
Researcher-Researcher	Via CRC/ Road-map meetings
	Meetings of the Streams for Dependencies
	Loose-coupling & Leveraging use of APIs
	Via Engineering Team

4.3.1. Role of Onsite Subject Matter Expert (SME).

We found that engaging the SME in the program has been significantly helpful to keep and monitor the users' perspectives in the innovation process. The SME facilitated the process by injecting the users' needs to the long-term as well as to the short-term visions of the program. The SME helped the program to scope the problem domain to apply research. Clarification of the problem domain at the initial phases enabled the researchers to produce an early prototype, and build trust with the users in terms of delivery. It was explained that: *"What was really great about this program is that [the SME] began to understand the research possibilities of the long-term vision, but also to a short-term need. Users have a really short-term need for solution so [building the early prototype] helped us to gain... The SME saw what we were doing, could match very closely to what they needed"*- ***P1- CRC Researcher/Manager.*** The SME supported the researchers and the engineers to understand users' needs by brokering links for end users and facilitated ad-hoc workshops to collect domain knowledge and feedback on the provided solutions. We argue that having this role was helpful for the program, given the complexity of law enforcement domain, the variety of users and the difficulties to access the confidential data.

4.3.2. Roadmap Meetings.

One of the frequently occurring themes in our analysis was the stakeholders' participation in 3-monthly roadmap meetings. The meetings were participated by the research streams, the CRC researchers and the engineers, the SME and any potential end-user. The meetings were organized for each stream separately, but cross-streams combined meetings could happen when required. The meetings were used as a platform to present the status of research, demonstrating any available output, coordinating the direction and collecting feedback from the involved parties. It is evident that the roadmaps intended to define and evaluate program milestones in a semi-formal way. One of the research leaders explained: *"the roadmap mentality was very good to turn a large course framework into what we're trying to do. But, we needed to [also] have, more intermediate meetings to refine that."* ***P3-Leader.*** The researchers highlighted that these meetings were helpful to collect feedback on research outputs and ensure alignment with users' needs. The roadmap meetings were also an attempt to simulate iterative/agile practices on research, yet with longer cycles (i.e., 3 months). The meetings were helpful to define deadlines for research outputs and enable incremental delivery.

4.3.3. Engineers Allocated to Research Streams.

Each research stream had a member from the engineering team assigned to it. The dedicated member was responsible to participate in weekly meetings with the corresponding research group for seeking and sharing updates. The engineers further discuss updates of research streams among themselves (e.g., in stand-up meetings). We found that this practice was helpful to coordinate between researchers and developers and facilitating integration of research outputs into end product. The engineers used these meetings to know the status of research and ensure alignment of the direction with users' needs and the product. They used the opportunity to obtain the knowledge about the ideas and technologies adopted at research side. The engineers highlighted that the weekly meetings enabled them to be aware of the mature outputs for integration and proactively start the process. One engineer stated: *"I get an understanding of the maturity about the technology and see if they do a demo. Sometimes I do the demo. So, if we see it's reached a certain level of maturity that might be useful to plug-in to our system, I would flag that to our (engineering) team here and, say, it seems they've got something that we could use or demonstrate to the customer."*- ***P9, Engineer***. From researchers' perspective the dedication practice helped to ensure the relevance of the outputs for CRC and the users. They used the practice to seek updates about the product and users' insights from engineers. One of the researchers highlighted that: *"It's definitely helpful. We understand what direction [CRC] is heading. It also confirms that our research is relevant to them and useful to them, which is motivating."* -***P11, Researcher***. We also found this practice helped resolve technological blockages that researchers could face due to dependencies to other streams. The engineers could suggest effective solutions to resolve dependencies between the streams, by taking over the issue in integration.

4.3.4. Collocation of Researchers with Engineers.

The researchers were encouraged to work at the engineers' location (e.g., 2-3 days a week). We observed that some of the researchers have been regularly visiting the CRC and had good experience. Our analysis revealed that collocation with engineers enabled researchers to gain technological and domain knowledge, which were helpful for their progress. All the visiting researchers had implementation tasks; hence, the engineers could support them to solve their development problems faster. One of the researchers stated his main motivation to visit the CRC as: *"because they can help me to solve the problems. Here, everybody is experienced in coding."*- ***P14, Researcher***. Beside the problem solving, it appeared that collocation was effective for technical coordination between the researchers and the engineers to integrate research outputs to the product. It was stated that: *"There are aspects of solution architecture that we don't want to deal with [...]. That communication is enormously facilitated by everybody, putting [them] in the same room a couple of times per week. I think both socially and technically the team benefits."* – ***P4, Leader.*** This practice enabled the researchers to better understand users' needs by: participating in ad-hoc meetings with users/SME when they come to the CRC, observing demo of the product, and gaining real scenarios from the engineers to apply their theoretical solutions. As of the frequent interaction with the SME and users, the engineers had developed a broader understanding of users' requirements to share with the researchers.

4.4. Challenges

Our analysis also revealed a number of challenges that the involved stakeholders faced in this collaboration model.

4.4.1. Users' Needs: Coping with Changes, Scoping Research.

It is well known that involving users in software development could create challenges for software team to cope with changes. We also found this issue emerging as a challenge in our data analysis. Our findings, however, demonstrate that coping with users' changes could be aggravated in science-based software development, as small changes in users' needs could have larger impact on the research directions and create frustrations. It was described: *"When you are changing direction, even if it's 5 degrees, there is a level of discomfort. In the academic sector, you're trained to be very deep in a very narrow area because you have to be world expert in that, and, that culture means people don't cope well in 5 degrees shift in their topic."- **P6, CRC Researcher/Manager***. We observed that this challenge was partly due to a lack of structured approach to involving users in the meetings. At the earlier phases of the program, the customer was sending any available resource (e.g., an investigator) to participate in meetings with the CRC/ researchers as of their commitment to the program. Yet, different individuals had different opinions about the problem domain. It is evident that having progressed in the program and increased maturity of the SME role, the CRC managed to fix this issue by involving a fixed number of users. The research leaders, on the other hand, alleviated the impact of user changes on their team by saving the PhD students at a backline. While Postdocs and research fellows were directly in charge of the program deliverables as *front-liners*, the PhD-students collaborated indirectly on the program.

Beside the changes in needs, it was highlighted that aligning users' needs (e.g., in form of product) with research ideas has been a great challenge, e.g., it was stated: "*We are developing [a product] that's going to be delivered out to use and going well. [Also] we have some research streams coming up with great ideas but matching those two together will be a challenge" – **P1, CRC Researcher/Manager.*** In this regard, some of the researchers highlighted the difficulty of placing the added value of research for users' problems, as the users were more interested in quick solutions. It was reported: "*Users don't see why this sophisticated algorithm is needed in their case. But, the purpose of the CRC is to look at how to improve end-user processes or systems, by injecting research capabilities." – **P16, Leader.*** Our analysis demonstrated that due to collaboration on applied research with user involvement, sometimes researchers had to compromise on their side and scope the direction to the users' needs. One of the researchers stated: *"you start reducing scope and eventually research questions purely to appease a specific end-user, then you are potentially cutting-off very fruitful terms of the trade from a research perspective"* ***– P10, Researcher.*** We observed that addressing users' needs also required researchers to prioritize their implementation tasks over pure research activities (e.g., publication). Whilst some researchers preferred development by nature, some others described it as a burden. We should note that the CRC was not purely interested in product development but had also KPIs defined for research publications. Nevertheless, some researchers stated that they needed more focus for publication, hence, tried to deliver the program deliverables first.

4.4.2. Difficulties of Collecting Proper User Feedback.

Our findings show that seeking users' feedback has been a challenging endeavor for researchers. Almost all the researchers underlined that users have very limited time, which creates difficulties in communicating with them in a timely manner. In addition, it was highlighted that although having an SME was helpful to connect to users, communication with them required researchers to use the users' language and culture. Some researchers found it difficult to make the users understand the areas they required feedback, without having them distracted on the other aspects of a prototype. As stated by an interviewee: *"I think one of the challenges of working with the end users is making sure that they understand what our research goals are and getting them to give feedback on our research goals, not on specifically what they are seeing in front of them."* – ***P11, Researcher***. It is not surprising that users could easily perceive and relate to the intuitive front-end solutions to share feedback. Yet, when it comes to more abstract, conceptual solutions at lower levels (e.g., architectural solution), the researchers faced more challenges. It was stated that: *"If I say I'm giving that complex capability, it's abstract. You can only see it once it actually exists and it is clearly demonstrable. Arguing in the absence of an existing implementation is more difficult."* – ***P4, Leader.*** While some of the research ideas required certain dataset for experimentation to progress, due to communication difficulties, the researchers' need was not understood and addressed at the side of users/customers. Furthermore, it is evident that the abstract, lower-level aspects of the system and the corresponding research work could be lower priority for some users. Accordingly, it was becoming challenging to get them engaged when not much interested in. In this regard, some of the researchers reflected that they had better experience communicating the added value of research, when they managed to talk to the right users. As stated:" *It's very hard. But the information management team has some interesting things. We call it 'duplication' for him. Because they know that when there is more duplication [in data] you have more noise. So, they want that problem to be solved."-**P5, Leader.***

4.4.3. Limitation in Access to Users' Data.

The researchers needed to have access to users' data in order to provide analytical solutions. Nevertheless, given the domain being national security, the required data most of the times was highly confidential and could not be shared. While CRC setup was successful to build connection between users and researchers, it could not do much in regard to the data access limitations. We observed that customers provided simulated datasets with anonymized/fake records to be used for research. Some of the researchers highlighted that though the data is simulated, it is still valuable as it presents real investigation scenarios. Some others stated that the sample data could support them only to a certain extent, or it is not ideal for their goals. One of the researchers described: "*The main difficulty, obviously in this particular domain is the ability to get access to the data that they are using. So, at this point, we have still very narrowly circumscribed set of data, which we adapt to demonstrate certain things."- **P4, Leader.***

4.4.4. Researchers – Engineers Collaboration Challenge. We described in section 4.2, the implementation of practices at CRC to promote collaboration between the researchers and the engineering team. Our findings, however, revealed several challenges that the two parties faced in this regard. We found that different priorities and viewpoints of the involved parties could cause tension. Whilst the engineers were concerned about the product and its quality, the researchers' goal was a proof of concept. The engineers believed that the code of the researchers is sometimes *"awful"*. Yet, a researcher stated: *"they look more into the quality of code, where we look into [our] goals. For an experiment, you don't need to bear to pass [variable] names, it's not a product."- **P13, Researcher.*** It was discussed by both sides that their goals sometimes conflict, and finding alignment is challenging. Some of the researchers reflected that this challenge is not specific to the engineering team per se but collaborating in a user-centric program like the CRC. They highlighted that they need to continuously do self-regulation to keep the balance between product and research work and avoid *"the risk of being too much caught up on the stakeholders and what they want" – **P10, Researcher.*** Our findings revealed that the researchers and the engineers faced communication gap due to cultural difference, language barriers, having different knowledge areas, and problem solving approaches. The research groups were mostly international coming from various countries and cultures. Apart from linguistic challenge, some of the engineers highlighted that they perceive a hierarchical culture in some of the groups that prohibits them to directly work with a researcher but going through the leader. Furthermore, it was discussed that researchers tend to be very narrow in their dialogue without sharing the context, which creates challenge for engineers to follow up.

5. LIMITATIONS

We discuss the potential validity threats that are relevant to this case study such as construct validity, external validity, and reliability according to the guidelines in [27, 28].

Construct validity questions the operational measures that have been employed to decrease the impact of subjective judgments [27]. Triangulation is one of the useful tactics in this regard [27, 28] that incorporates different angles of the phenomenon under study into consideration. We involved at least one interviewee from each stakeholder group to cover a comprehensive view. We talked with different roles e.g., managers, leaders, researchers and developers from the CRC. The majority of the interviews were conducted by two researchers in order to complement the discussion points and remaining on the topic given the open-ended nature of the questions. We tried to address any doubt in data analysis through internal discussions. Furthermore, we presented the results to the CRC to validate our understanding. Yet, given the nature of qualitative data, we confirm that the results are based on our understanding and interpretation.

External validity questions generalizability of findings. Case study reports findings based on analysis of a phenomenon in its real-world context. It varies from research methods that use statically representative samples taken from a population to produce results [27, 28]. In the context of case study, the objective is to ensure analytical generalization of findings to the cases with similar context [27]. In this paper, we present findings based on a single case study in the context of CRC collaboration models. Our results could be only applicable to a similar context (e.g., CRC setup, law enforcement domain).

Reliability questions the extent that results are dependent on a researcher [27, 28]. We addressed this threat by maintaining proper documentation for all the steps of this study. We prepared research protocol and interview guidelines[2] (i.e., customized based on role of interviewees). All the interviews were audio recorded and verbatim transcribed for the sake of refer-ability and reliability. We performed the analysis using Nvivo. It supported us to navigate between emerging codes from different data sources and systematically perform comparison, revise/ merge the codes. The extracted data and our coding scheme are recorded in the tool.

6. DISCUSSION

In the followings, we present the main takeaways from our study and discuss the implications for research and practice.

6.1. Takeaways from Findings

The CRC User-Centric Approach to Software Innovation. Existing literature present different process models describing creation and transfer of software technology innovation in collaboration with academia with high-tech sector (e.g., [11], [2]). We investigated the processes and practices of one of the Australian CRCs, D2D, for establishing and sustaining collaboration among researchers, software engineers, and end-users to provide innovative software technologies. We described how the CRC has facilitated

[2] Available at: https://drive.google.com/drive/folders/13OmS198d8dQGC-eBZOBPlfVV0CTTVVMp

engagements by establishing trust among the key stakeholders. We argue that the implemented approach of the CRC presented here differs from several other R&D models that take place in direct collaboration of academics with large companies (e.g., ABB and Philips). As in those cases typically innovation takes place in an organizational context (e.g., for improving current product, process)[11], [2]. The D2D's software product itself was considered innovative. There was no sophisticated solution based on scientific research to leverage big data for the Australian National security and law enforcement domain. Given the complexity of the domain and the scarcity of a similar product, it was crucial to involve users to extract the domain knowledge and ensure the relevance of the invented solutions to the users' needs.

Involving users in new product development has been practiced for several years. Users could be involved in part of the process such as early phases of innovation (e.g., via focus group) [29], market analysis and validation [15], or through the whole process systematically [19]. Here, the users were engaged in the project initiation and the execution phases. The CRC's setup enabled the researchers to interact with users, understand the real-world scenarios and access the users' data from a highly confidential and complex business domain. The academics believed in the *applied science*, and valued the user-centric cooperation model. The researchers could see the impact of their research through close collaboration with end-users in the CRC consortium. It was crucial to involve the academic researchers that had the positive attitude towards the applied science to successfully implement this model. It was underlined that there were incidents that D2D had to cut the funding and collaboration with academics, as they were not willing to align research directions with the product priorities. Innovation in its definition emphasizes on diffusion of the invention to end-users [9], yet it is evident that often the outputs of research projects are not mature enough for transfer [12]. We argue that the CRC's user-centric approach of software innovation helps centralize coordination to facilitate information flow among stakeholders. It ensures relevance of scientific outputs for end-users and enables product delivery through its setup (e.g., dedicating engineering team). It is, however, very important to implement suitable practices to address the diverse needs and interests of all the involved parties and keep them motivated and engaged during the R&D process.

Agility in the Software Innovation Process. In the context of software development user involvement has been promoted through different techniques, in particular agile methods [16]. We also observed that adopting a user-centric approach for transferring science to technology enforced *agility* in the innovation process. The user engagement was challenging since it required building trust with users and motivating them to invest time and share knowledge. Hence, moving towards agility in this case was a strategy to facilitate user engagement. Agility in the context of our study was manifested in: incremental delivery of viable product, involving customers/users in the process (e.g., SME as product owner) and tolerating (some) changes. Whilst the engineers used Scrum to enhance the product baseline, the researchers followed a longer cycle (i.e., 3 months) to share updates and deliverables. We observed that user-focus was retained in the process using two strategies: a) maintaining regular interaction of engineers and researchers with the SME (and end-users when required), b) making close ties between engineers and researchers. We found that engineers also acted as an interface to translate users' needs for researchers. Compared with researchers, they had better communication skills and more flexibility to interact with users.

Our study is one of the few studies that have empirically explored the use of agile methodology in academia-industry collaboration. Sandberg and Crnkovic [5] present successful use of Scrum in direct collaboration of researchers with industrial partners. They found that Scrum practices were helpful to build trust, and facilitate knowledge sharing and technology transfer [5]. Yet, in their study the teams were collaborating on separate research projects, not a common product that we observed. We found that agility and empowering users could also introduce challenges for researchers in terms of getting too much influenced by users, scoping research and difficulties in seeking proper inputs. We observed that this CRC has tried to alleviate some of these challenges by practices: e.g., value publications, have longer cycles for research, contractual flexibility to accommodate changes, decreasing development load from researchers by involving engineers. However, we believe that more in-depth exploration is required to further improve the process of user-centric software innovation in the CRC programs.

6.2. Implications

We discuss the implications of our study for research and practice in the following subsections.

Research. The availability of big data has largely raised demands for sophisticated software solutions using scientific data analytical skills. This opportunity has introduced a new trend in software technology innovation to leverage big data problems. We presented how such innovation process could be implemented in practice using the CRC collaboration model, and described the associated challenges. To the best of our knowledge, this work is among the first studies that empirically investigate software innovation process for providing data-driven solutions using CRC model. This complicated collaboration model requires more empirical investigation to realize the stakeholders' challenges and provide best practices. The future research could investigate the practices to effectively balance the expectations and KPIs of the involved stakeholders in the process. Devising mechanisms for efficient coordination is also another interesting area to explore further. Such exploration could uncover the optimal team size and trade-offs between involved capability areas vs. the coordination effort.

Practice. Our study provides evidence-based knowledge about the processes and practices of enabling tight academia-industry collaboration for rapidly transferring scientific knowledge of data science to innovative software solutions. It highlights the potentials to engage data analytical science to apply on the increasing demands in the market. The CRC collaboration model, though being implemented in some countries for several years, could set example for practitioners in other geographies that the model is not common. In addition, we believe a potential comparison of software innovation process in CRC-based vs. organizational-based (i.e., direct collaboration of industry with universities) model could be beneficial to understand pros and cons of each approach and adopt practices from one setting to another.

7. CONCLUSION & FUTURE WORK

We have presented the findings from a case study in the context of a CRC program aimed at enabling collaboration between academia, industry and end-users for developing data-driven science and software technologies. We described the innovation process implemented to rapidly develop fundamental data science approaches and leverage them to build and deploy software solutions for the end users in the Australian law enforcement domain. We have elaborated the coordination mechanisms and the associated challenges of implementing a user-centric approach of software technology innovation. We took the user-involvement angle to elaborate the observed practices and challenges. Nevertheless, our data has revealed other interesting themes in relation to coordination among research streams and the challenges of transferring research outputs to software products. We intend to explore this angle in our future work. Furthermore, we also plan to conduct more case studies with other CRCs [6] in Australia. This step will help us to enrich and complement the current findings towards generalization.

ACKNOWLEDGMENTS

We greatly appreciate the time and valuable insights of our participants from CRC as well as research groups.

REFERENCES

[1] A. Sandberg, L. Pareto, and T. Arts, "Agile collaborative research: Action principles for industry-academia collaboration," *IEEE software,* vol. 28, pp. 74-83, 2011.

[2] T. Gorschek, P. Garre, S. Larsson, and C. Wohlin, "A model for technology transfer in practice," *IEEE software,* vol. 23, pp. 88-95, 2006.

[3] P. Runeson and S. Minör, "The 4+ 1 view model of industry--academia collaboration," in *Proceedings of the 2014 international workshop on Long-term industrial collaboration on software engineering*, 2014, pp. 21-24.

[4] V. Garousi, K. Petersen, and B. Ozkan, "Challenges and best practices in industry-academia collaborations in software engineering: A systematic literature review," *Information and Software Technology,* vol. 79, pp. 106-127, 2016.

[5] A. B. Sandberg and I. Crnkovic, "Meeting industry-Academia research collaboration challenges with agile methodologies," in *Proceedings of the 39th International Conference on Software Engineering: Software Engineering in Practice Track*, 2017, pp. 73-82.

[6] (20th July 2018). *Cooperative Research Centres (CRC) Program*. Available: https://crca.asn.au

[7] E. Sinnewe, M. B. Charles, and R. Keast, "Australia's Cooperative Research Centre Program: A transaction cost theory perspective," *Research Policy,* vol. 45, pp. 195-204, 2016.

[8] Z. J. Acs and D. B. Audretsch, "Innovation in large and small firms: an empirical analysis," *The American economic review,* pp. 678-690, 1988.

[9] R. Garcia and R. Calantone, "A critical look at technological innovation typology and innovativeness terminology: a literature review," *Journal of Product Innovation Management: AN INTERNATIONAL PUBLICATION OF THE PRODUCT DEVELOPMENT & MANAGEMENT ASSOCIATION,* vol. 19, pp. 110-132, 2002.

[10] H. Edison, N. Bin Ali, and R. Torkar, "Towards innovation measurement in the software industry," *Journal of Systems and Software,* vol. 86, pp. 1390-1407, 2013.

[11] T. Punter, R. L. Krikhaar, and R. J. Bril, "Software engineering technology innovation–Turning research results into industrial success," *Journal of Systems and Software,* vol. 82, pp. 993-1003, 2009.

[12] P. Diebold, A. Vetro, and D. M. Fernández, "An exploratory study on technology transfer in software engineering," in *Empirical Software Engineering and Measurement (ESEM), 2015 ACM/IEEE International Symposium on*, 2015, pp. 1-10.

[13] H. Munir, K. Wnuk, and P. Runeson, "Open innovation in software engineering: a systematic mapping study," *Empirical Software Engineering,* vol. 21, pp. 684-723, 2016.

[14] W. Vanhaverbeke, V. Van de Vrande, and H. Chesbrough, "Understanding the advantages of open innovation practices in corporate venturing in terms of real options," *Creativity and innovation management,* vol. 17, pp. 251-258, 2008.

[15] H. H. Olsson and J. Bosch, "Collaborative Innovation: A Model for Selecting the Optimal Ecosystem Innovation Strategy," in *Software Engineering and Advanced Applications (SEAA), 2016 42th Euromicro Conference on*, 2016, pp. 206-213.

[16] S. Chamberlain, H. Sharp, and N. Maiden, "Towards a framework for integrating agile development and user-centred design," in *International Conference on Extreme Programming and Agile Processes in Software Engineering*, 2006, pp. 143-153.

[17] D. G. Feitelson, E. Frachtenberg, and K. L. Beck, "Development and deployment at facebook," *IEEE Internet Computing,* vol. 17, pp. 8-17, 2013.

[18] G. Schermann, J. Cito, P. Leitner, U. Zdun, and H. C. Gall, "We're doing it live: A multi-method empirical study on continuous experimentation," *Information and Software Technology,* 2018.

[19] P. Bosch-Sijtsema and J. Bosch, "User involvement throughout the innovation process in high-tech industries," *Journal of Product Innovation Management,* vol. 32, pp. 793-807, 2015.

[20] D. Rombach and R. Achatz, "Research collaborations between academia and industry," in *2007 Future of Software Engineering*, 2007, pp. 29-36.

[21] E. Villani, E. Rasmussen, and R. Grimaldi, "How intermediary organizations facilitate university–industry technology transfer: A proximity approach," *Technological Forecasting and Social Change,* vol. 114, pp. 86-102, 2017.

[22] E. M. Mora-Valentin, A. Montoro-Sanchez, and L. A. Guerras-Martin, "Determining factors in the success of R&D cooperative agreements between firms and research organizations," *Research Policy,* vol. 33, pp. 17-40, 2004.

[23] C. Buller and W. Taylor, "Partnerships between public and private: The experience of the cooperative research center for plant science in Australia," 1999.

[24] S. Garrett-Jones, T. Turpin, and K. Diment, "Careers and organisational objectives: managing competing interests in cooperative research centres," in *Cooperative Research Centers and Technical Innovation*, ed: Springer, 2013, pp. 79-110.

[25] R. Yin, *Case study research: Design and methods*: Sage Publications, Inc, 2003.

[26] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology,* vol. 3, 2006.

[27] R. K. Yin, *Applications of case study research*, 3rd ed.: SAGE Publications, Incorporated, 2012.

[28] P. Runeson and M. Host, "Guidelines for conducting and reporting case study research in software engineering," Empirical Software *Engineering,* vol. 14, pp. 131-164, 2009/04/01 2009.

[29] E. Van Kleef, H. C. Van Trijp, and P. Luning, "Consumer research in the early stages of new product development: a critical review of methods and techniques," *Food quality and preference,* vol. 16, pp. 181-201, 2005.